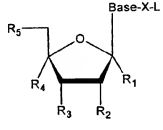
Macromolecular nucleotide compounds and methods of their application

Claims:

1. Macromolecular compounds with the structure:
$(nuc\text{-}linker)_n$-marker

wherein:
Nuc is a nucleotide or nucleoside (nuc-component)
Linker is a linker component comprising the following parts:
a) Coupling unit L is a part of the linker which provides the linkage between nuc and the rest of the linker
b) Polymer is a part of the linker which is a water-soluble polymer with an average length between 100 and 20,000 atoms (chain atoms)
c) Coupling unit T is a part of the linker which provides the linkage between the marker and the rest of the linker
Marker is a marker component
(n) is a positive integer between 1 and 100

2. Macromolecular compounds according to claim 1, wherein the nuc-component comprises the following structures:



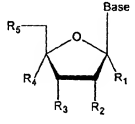
Wherein:
Base is selected independently from the group of adenine, or 7-deazaadenine, or guanine, or 7-deazaguanine, or thymine, or cytosine, or uracil, or their modifications, wherein X is the coupling position of the linker to the base and L is the coupling unit of the linker (L).
R_1 - is H
R_2 - is selected independently from the group of H, OH, halogen, NH_2, SH or protected OH group

R_3 - is selected independently from the group of H, OH, halogen, PO_3, SH, N_3, NH_2, O-CH_3, O-CH_2-O-CH_3, O-CH_2-CH=CH_2, O-R_{3-1}, $P(O)_m$-R_{3-1} ((m) is 1 or 2), NH-R_{3-1}, S-R_{3-1}, Si-R_{3-1} wherein R_{3-1} is a chemically, photochemically or enzymatically cleavable group.

R_4 - is H or OH

R_5 - is selected independently from the group of OH, or a protected OH group, or a monophosphate group, or a diphosphate group, or a triphosphate group, or is an alpha thiotriphosphate group.

3. Macromolecular compounds according to claim 1, wherein the nuc-component comprises the following structures:



Wherein:

Base is selected independently from the group of adenine, or 7-deazaadenine, or guanine, or 7-deazaguanine, or thymine, or cytosine, or uracil, or their modifications capable of enzymatic reactions.

R_1 - is H

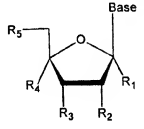
R_2 - is selected independently from the group of H, OH, halogen, NH_2, SH or protected OH group

R_3 - is selected independently from the group of O- R_{3-2}-L, $P(O)_m$- R_{3-2}-L and (m) is 1 or 2, NH-R_{3-2}-L, S-R_{3-2}-L, SI-R_{3-2}-L, wherein R_{3-2} is the coupling position of the linker to the nucleotide and L is the coupling unit of the linker (L).

R_4 - is H or OH

R_5 - is selected independently from the group of OH, or a protected OH group, or a monophosphate group, or a diphosphate group, or a triphosphate group, or is an alpha-thiotriphosphate group.

4. Macromolecular compounds according to claim 1, wherein the nuc-component comprises the following structures:



Wherein:

Base is selected independently from the group of adenine, or 7-deazaadenine, or guanine, or 7-deazaguanine, or thymine, or cytosine, or uracil, or their modifications capable of enzymatic reactions.

R_1 - is H

R_2 - is selected independently from the group of H, OH, halogen, NH_2, SH or protected OH group

R_3 - is selected independently from the group of H, OH, halogen, PO_3, SH, NH_2, $O\text{-}R_{3\text{-}1}$, $P(O)_m\text{-}R_{3\text{-}1}$ ((m) is 1 or 2), $NH\text{-}R_{3\text{-}1}$, $S\text{-}R_{3\text{-}1}$, $Si\text{-}R_{3\text{-}1}$ wherein $R_{3\text{-}1}$ is a chemically, photochemically or enzymatically cleavable group.

R_4 - is H or OH

R_5 - is selected independently from the group of $O\text{-}R_{5\text{-}1}\text{-}L$, or $P\text{-}(O)_3\text{-}R_{5\text{-}1}\text{-}L$ (modified monophosphate group), or $P\text{-}(O)_3\text{-}P\text{-}(O)_3\text{-}R_{5\text{-}1}\text{-}L$ (modified diphosphate group) or $P\text{-}(O)_3\text{-}P\text{-}(O)_3\text{-}P\text{-}(O)_3\text{-}R_{5\text{-}1}\text{-}L$ (modified triphosphate group), wherein $R_{5\text{-}1}$ is the coupling position of the linker to the nucleotide and L is the coupling unit of the linker (L).

5. Macromolecular compounds according to claims 1 to 4, wherein the coupling unit (L) of the linker comprises the following structural elements:

$R_6\text{-}NH\text{-}R_7$, $R_6\text{-}O\text{-}R_7$, $R_6\text{-}S\text{-}R_7$, $R_6\text{-}SS\text{-}R_7$, $R_6\text{-}CO\text{-}NH\text{-}R_7$, $R_6\text{-}NH\text{-}CO\text{-}R_7$, $R_6\text{-}CO\text{-}O\text{-}R_7$, $R_6\text{-}O\text{-}CO\text{-}R_7$, $R_6\text{-}CO\text{-}S\text{-}R_7$, $R_6\text{-}S\text{-}CO\text{-}R_7$, $R_6\text{-}P(O)_2\text{-}R_7$, $R_6\text{-}Si\text{-}R_7$, $R_6\text{-}(CH_2)_n\text{-}R_7$, $R_6\text{-}(CH_2)_n\text{-}R_7$, $R_6\text{-}A\text{-}(CH_2)_n\text{-}R_7$, $R_6\text{-}(CH_2)_n\text{-}B\text{-}R_7$, $R_6\text{-}(CH{=}CH\text{-})_n\text{-}R_7$, $R_6\text{-}(A\text{-}CH{=}CH\text{-})_n\text{-}R_7$, $R_6\text{-}(CH{=}CH\text{-}B\text{-})_n\text{-}R_7$, $R_6\text{-}A\text{-}CH{=}CH\text{-}(CH_2\text{-})_n\text{-}R_7$, $R_6\text{-}(\text{-}CH{=}CH\text{-}CH_2)_n\text{-}B\text{-}R_7$, $R_6\text{-}(C{\equiv}C\text{-})_n\text{-}R_7$, $R_6\text{-}(A\text{-}C{\equiv}C\text{-})_n\text{-}R_7$, $R_6\text{-}(C{\equiv}C\text{-}B\text{-})_n\text{-}R_7$, $R_6\text{-}A\text{-}C{\equiv}C\text{-}(CH_2\text{-})_n\text{-}R_7$, $R_6\text{-}(\text{-}C{\equiv}C\text{-}CH_2)_n\text{-}B\text{-}R_7$, $R_6\text{-}(\text{-}C{\equiv}C\text{-}CH_2\text{-}CH_2\text{-})_n\text{-}B\text{-}R_7$,

wherein R_6 is the nuc-component, R_7 is the rest of the linker, and A and B comprise the following structural elements: -NH- , -O- , -S- , -SS- , -CO-NH- , -NH-CO- , -CO-O- , -O-CO- , -CO-S- , -S-CO- , $-P(O)_2-$, -Si-, $-(CH_2)_n-$, wherein (n) ranges from 1 to 5, a photolabile group

6. Macromolecular compounds according to claims 1 to 5, wherein the linker-component comprises a water-soluble polymer.

7. Macromolecular compounds according to claim 6, wherein the linker-component comprises water-soluble polymers selected independently from the following group:
polyethylene glycol (PEG), polysaccharides, dextran, polyamides, polypeptides, polyphosphates, polyacetates, polyalkyleneglycoles, copolymers from ethyleneglycol and propyleneglycol, polyolefinic alcohols, polyvinylpyrrolidones, poly(hydroxyalkylmethacrylamides), polyhydroxyalkylmethacrylates, poly(x-hydroxy) acids, polyacrylic acid, polyacrylamide, polyvinylalcohol.

8. Macromolecular compounds according to claims 1 to 7, wherein the average length of a linker component ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 10000, 10000 to 50000 atoms (chain atoms).

9. Macromolecular compounds according to claims 1 to 8, wherein a marker component has one of the following functions: signal-giving function, signal-transmitting function, catalytic function or affine function.

10. Macromolecular compounds according to claims 1 to 9, wherein a marker component consists of one structural marker unit.

11. Macromolecular compounds according to claims 1 to 9, wherein a marker component consists of several structural marker units bonded to a core component.

12. Macromolecular compounds according to claims 10 or 11, wherein a structural marker unit independently comprises one of the following structural elements: biotin, hapten, radioactive isotope, rare-earth atom, dye, fluorescent dye.

13. Macromolecular compounds according to claims 10 or 11, wherein a structural marker unit independently comprises one of the following elements: nanocrystals or their modifications, proteins or their modifications, nucleic acids or their modifications, particles or their modifications.

14. Macromolecular compounds according to claim 13, wherein a structural marker unit comprises one of the following proteins:
enzymes or their conjugates or modifications,
antibodies or their conjugates or modifications,
streptavidin or its conjugates or modifications,
avidin or its conjugates or modifications

15. Macromolecular compounds according to claim 13, wherein a structural marker unit comprises one of the following types of nucleic acid chains: DNA, RNA, PNA, wherein the length of nucleic acid chains ranges between 10 and 10,000 nucleotides.

16. Macromolecular compounds according to claims 11 to 15, wherein the core component of the marker component independently comprises one of the following elements: water-soluble Polymer from the group of: polyamides (e.g. polypeptides), polyacrylic acid and its derivates, polyacrylamides and their derivates, polyvinylalcohols and their derivates, nucleic acids and their derivates, streptavidin or avidin and their derivates, dendrimeres, whereas these elements can be linear or branched or crosslinked with each other.

17. Macromolecular compounds according to claims 1 to 9 and 11 to 16, wherein the linkage between several structural marker units and the core component is covalent or affine.

18. Macromolecular compounds according to claims 1 to 10, wherein the linkage between a structural marker unit and the linker is covalent or affine.

19. Macromolecular compounds according to claims 1 to 9, 11 to 17, wherein the linkage between the core component and the linker is covalent or affine.

20. Macromolecular compounds according to claims 1 to 19, wherein only one nuc-component with one linker component is linked to the marker component, wherein the linker length ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms.

21. Macromolecular compounds according to claims 1 to 20, wherein only one nuc-component with one linker component is linked to the marker component, wherein the linker length ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms and the linker component comprises one or several compounds that are cleavable under mild conditions.

22. Macromolecular compounds according to claims 1 to 21, wherein only one nuc-component with one linker component is linked to the marker component, wherein the linker length ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms and one or several parts of the nuc-macromolecule are modified in such a way, that only one nuc-component can be incorporated into the growing strand.

23. Macromolecular compounds according to claims 1 to 19, wherein several nuc-components are each coupled to one marker component via a linker, wherein the length of each respective linker component ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms.

24. Macromolecular compounds according to claims 1 to 19, 23, wherein several nuc-components are each coupled to one marker component via a linker, wherein the length of each respective linker component ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms and each respective linker component comprises one or several compounds that are cleavable under mild conditions.

25. Macromolecular compounds according to claims 1 to 19, 23 and 24, wherein several nuc-components are each coupled to one marker component via a linker, wherein the length of each respective linker component ranges between 50 to 100, 100 to 200, 200 to 500, 500 to 1000, 1000 to 2000, 2000 to 5000 atoms, and one or several parts of the nuc-macromolecule are modified in such a way that only one nuc-component can be incorporated into the growing nucleic acid chain.

26. Oligonucleotides or polynucleotides comprising at least one nuc-macromolecule according to claims 1 to 25 per one nucleic acid chain.

27. Oligonucleotides or polynucleotides according to claim 26, wherein oligo- or polynucleotides are RNA or DNA or PNA and their length ranges between 5 and 50,000 nucleotides.

28. Method of modification of nucleic acid chains, wherein nuc-macromolecules according to claims 1 to 25 are used for the coupling.

29. Method according to the claim 28, wherein the modification is accomplished by an enzymatic coupling and the reaction mixture comprises the following components:

- at least one type of nuc-macromolecules or their intermediate stages according to the claims 1 to 25, wherein every type of nuc-macromolecule is distinctively labeled,
- at least one population of the nucleic acid chains,
- at least one type of enzyme for coupling nuc-macromolecules to the nucleic acid chains,

30. Method according to claim 28, wherein the modification is accomplished by an enzymatic coupling and the reaction mixture comprises the following components:

- at least one type of nuc-macromolecules or their intermediate stages according to the claims 1 to 25, wherein every type of nuc-macromolecule is distinctively labeled,
- at least one population of the nucleic acid chains,
- at least one type of enzyme for coupling nuc-macromolecules to the nucleic acid chains,
- at least one other type of nucleoside triphosphates.

31. Method according to claims 29, 30, wherein the said type of enzyme independently comprises one of the following groups: DNA-polymerases, RNA-polymerases, terminal transferases.

32. Method according to claim 30, wherein the "other type" of nucleoside triphosphates is independently selected from the group of ribonucleoside triphosphates (ATP, GTP, UTP, CTP), of 2`-deoxyribonucleoside triphosphates (dATP, dUTP, dTTP, dCTP, dGTP), of 2`,3`-dideoxynucleoside triphosphates (ddATP, ddGTP, ddUTP, ddCTP, ddTTP).

33. Method according to claim 32, wherein the "other type" of nucleoside triphosphates is conventionally modified nucleotides with a label, wherein the said label is independently selected from the group of fluorescent dye, biotin, hapten or radioactive element.

34. Method according to claims 28 to 33, wherein at least two different populations of nucleic acid chains are present

35. Method according to claim 34, wherein at least one of the populations of the nucleic acid chains has a primer function and at least one population of the nucleic acid chains has a template function.

36. Method according to claim 28, wherein the modification is accomplished by chemical coupling and the coupling of the nuc-macromolecules to the nucleic acid chain is accomplished via phosphoroamidite-coupling.

37. Method according to claims 28 to 36, wherein nuc-macromolecules which allow for the coupling of only single nuc-component into the growing nucleic acid strand are used for the labeling process and multiple incorporations are prevented by modifications of the nuc-component and/or the linker component and/or the marker component.

38. Method according to claim 37, wherein the multiple coupling is prevented reversibly.

39. Method according to claim 37, wherein the multiple coupling is prevented irreversibly.

40. Method according to claims 28 to 36, wherein nuc-macromolecules which allow for the coupling of multiple nuc-components into the growing nucleic acid strand are used for the labelling process.

41. Method according to claims 28 to 40, wherein the nucleic acid chains participating in the reaction are coupled to a solid phase and have addressable positions.

42. Method according to claim 41, wherein the nucleic acid chains compose a uniform population.

43. Method according to claim 41, wherein the nucleic acid chains compose two or more different populations and each of the populations has an addressable position on the solid phase.

44. Method according to claims 41, 42, wherein the coupling of nuc-macromolecules is conducted on the uniform population of nucleic acid molecules attached to the solid phase and the marker component of the nuc-macromolecule remains on the extended nucleic acid strand after the coupling and is not cleaved off.

45. Method according to claims 41, 42, wherein the coupling of the nuc-macromolecules is conducted on the uniform population of nucleic acid chains attached to the solid phase and the marker component or its individual parts are cleaved off, with or without the linker component of the nuc-macromolecule, from the nuc-component incorporated into the growing nucleic acid strand, the cleaving-off taking place during or after the coupling.

46. Method according to claims 41, 43, wherein the coupling of nuc-macromolecules in a reaction mixture is conducted simultaneously on two or more different populations of nucleic acid chains attached to the solid phase, wherein each of these populations has distinct addressable positions on the solid phase, and the marker component of the nuc-macromolecule remains on the extended nucleic acid strand after the coupling and is not cleaved off.

47. Method according to the claims 41, 43, wherein the coupling of nuc-macromolecules is conducted simultaneously on two or more different populations of nucleic acid chains attached to the solid phase, wherein each of these populations has distinct addressable positions on the solid phase, and the marker component or its individual parts are cleaved off, with or without linker component of the nuc-macromolecule, from the nuc-component, the cleaving-off taking place during or after the coupling.

48. Method according to claims 41 to 47, wherein the addressable positions having nucleic acid molecules on the solid phase are distributed as spots on a plane surface, and nucleic acid molecules are uniform on each spot.

49. Method according to the claims 41 to 47, wherein the addressable positions having nucleic acid molecules are fastened on the beads or particles and nucleic acid molecules are uniform for each bead.

50. Method according to claims 41 to 47, wherein the addressable positions having nucleic acid molecules are distributed in a multivessel array, like a microtiter plate or nanotiter plate or picotiter plate, wherein the nucleic acid molecules are uniform in one vessel of the mutlivessel array.

51. Method according to the claims 28 to 35 and 37 to 50, which comprises the following steps:

a) Providing of at least one population of single-stranded nucleic acid chains (NAC),
b) Hybridizing primers to these nucleic acid chains, whereas extendable NAC primer complexes are formed,
c) Incubation of at least one type of the nuc-macromolecule according to claims 1 to 25 together with a type of polymerase according to claim 31 with provided NAC primer complexes in steps (a) and (b) under conditions which allow for incorporation of complementary nuc-macromolecules, and each kind of the nuc-macromolecule having a distinctive label,
d) Removal of the unincorporated nuc-macromolecules from the NAC primer complexes,
e) Detection of the signals from the nuc-macromolecules which are incorporated in the NAC primer complexes,
f) Removal of the linker component and the marker component from the nuc-macromolecules which are incorporated in the NAC primer complexes,
g) Wash the NAC primer complexes,

if necessary, repetition of the steps (c) to (g).

52. Method according to the claims 28 to 40, wherein the nucleic acid chains are coupled to a solid phase in a random arrangement.

53. Method according to claims 28 to 41, 52 for the parallel sequence analysis of nucleic acid sequences (nucleic acid chains, NACs), in which

fragments (NACFs) of single-stranded NACs with a length of approximately 50 to 1000 nucleotides that may represent overlapping partial sequences of the whole sequence are produced,

the NACFs are bonded to a reaction surface in a random arrangement using a uniform or several different primers in the form of NACF primer complexes,

a cyclical synthesis reaction of the complementary strand of the NACFs is performed using one or more polymerases by

a) adding, to the NACF primer complexes bonded to the surface, a solution containing one or more polymerases and one to four nuc-macromolecules that have a marker component labeled with fluorescent dyes, wherein the fluorescent dyes, which each are located on the marker component when at least two nuc-macromolecules are used simultaneously, are chosen in such a manner that the nuc-macromolecules used can be distinguished from one another by measurement of different fluorescent signals, the nuc-macromolecules being structurally modified in such a manner that the polymerase is not capable of incorporating another nuc-macromolecule in the same strand after such a nuc-macromolecule has been incorporated in a growing complementary strand, the linker component and marker component being cleavable,

b) incubating the stationary phase obtained in step a) under conditions suitable for extending the complementary strands, the complementary strands each being extended by one nuc-macromolecule,

c) washing the stationary phase obtained in step b) under conditions suitable for removing nuc-macromolecules that are not incorporated in a complementary strand,

d) detecting the single nuc-macromolecules incorporated in complementary strands by measuring the characteristic signal of the respective fluorescent dye, the relative position of the individual fluorescent signals on the reaction surface being determined at the same time,

e) cleaving-off the linker component and marker component of the nuc-components added to the complementary strand in order to produce unlabeled (nucleotides or) NACFs,

f) washing the stationary phase obtained in step e) under conditions suitable for the removal of the marker component,

repeating steps a) to f), several times if necessary,

the relative position of individual NACF primer complexes on the reaction surface and the sequence of these NACFs being determined by specific assignment of the fluorescent signals, which were detected in the respective positions in step d) during successive cycles, to the nuc-macromolecules.

54. Method according to claim 53, characterized in that steps a) to f) of the cyclical synthesis reaction are repeated several times, only one type of nuc-macromolecule being used in each cycle.

55. Method according to claim 53 characterized in that steps a) to f) of the cyclical synthesis reaction are repeated several times, two types of differently labeled nuc-macromolecules being used in each cycle.

56. Method according to claim 53 characterized in that steps a) to f) of the cyclical synthesis reaction are repeated several times, four types of differently labeled nuc-macromolecules being used in each cycle.

57. Method according to claim 53 characterized in that the NACs are variants of a known reference sequence and steps a) to f) of the cyclical synthesis reaction are repeated several times, two differently labeled types of nuc-macromolecules and two unlabeled nucleotides being used alternately in the cycles and the whole sequences being determined by comparison with the reference sequence.

58. Method according to claims 53 to 57 characterized in that a primer binding site (PBS) is introduced in each of the NACFs, one PBS being introduced at both complementary single strands in the case of double-stranded NACs and the primer binding sites displaying identical or different sequences for all NACFs.

59. Method according to claims 53 to 57 characterized in that the NACFs are brought into contact with primers in a solution under conditions suitable for the hybridization of the primers to the primer binding sites (PBSs) of the NACFs, the primers exhibiting identical or different sequences to one another, and the NACF primer complexes formed then being bonded to the reaction surface.

60. Method according to claims 53 to 57 characterized in that the NACFs are first of all immobilized on the reaction surface and only then brought into contact with primers under conditions suitable for the hybridization of the primers to the primer binding sites (PBSs) of the NACFs, NACF primer complexes being formed, the primers exhibiting identical or different sequences to one another.

61. Method according to claims 53 to 60, wherein the incorporation reaction is being performed simultaneously on 10 to 100,000 different sequence populations.

62. Method according to claims 53 to 60, wherein the incorporation reaction is being performed simultaneously on 100,000 to 100,000,000 different sequence populations.

63. Method according to claims 28 to 62, wherein sequences of the nucleic acid chains are determined.

64. Method according claims 28 to 63, wherein the marker component is fluorescently labeled.

65. Method according claims 41 to 64, wherein the solid phase is independently selected from the following group: silicone, glass, ceramics, plastics, gels or their modifications.

66. A kit comprising the macromolecular compounds according to claims 1 to 25.

67. Methods using the macromolecular compounds according to claims 1 to 25 in enzymatic reactions, wherein enzymes comprise the following groups: polymerases, ligases, nucleases (endo or exonucleases).

68. Methods using the macromolecular compounds according to claims 26 and 27 in enzymatic reactions, wherein enzymes comprise the following groups: polymerases, ligases, nucleases (endo or exonucleases).